Ecological Indicators 85 (2018) 820–831

Contents lists available at ScienceDirect

Ecological Indicators

journal homepage: www.elsevier.com/locate/ecolind

Predicting the effect of land use and climate change on stream macroinvertebrates based on the linkage between structural equation modeling and bayesian network

Xue Li[a], Yuan Zhang[b,*], Fen Guo[b], Xin Gao[b], Yuqiu Wang[c,*]

[a] Tianjin Key Laboratory of Water Resources and Environment, Tianjin Normal University, Tianjin, 300387, China
[b] State Key Laboratory of Environmental Criteria and Risk Assessment and Laboratory of Riverine Ecological Conservation and Technology, Chinese Research Academy of Environmental Sciences, Beijing, 100012, China
[c] College of Environmental Science and Engineering, Nankai University, Tianjin, 300071, China

ARTICLE INFO

Keywords:
Freshwater ecosystems
Structural equation modeling
Bayesian networks
Human disturbances
Climate change

ABSTRACT

Land use and climate change are increasingly important stressors affecting freshwater ecosystems. Although their effects on freshwater organisms have been widely studied, most studies applied traditional statistical methods, which only focus on single stressor and simple cause-effect relationships without considering the complex interactions among stressors. Therefore, we developed an integrated method by combining Structure Equation Modeling (SEM) and Bayesian Networks (BNs) to estimate the interactive effect of land use and climate change on freshwater macroinvertebrates. A field investigation was conducted in August 2009 in Taizi River Basin, Northeast China, and samples of stream macroinverterbates and water chemistry were collected from 211 sites. The SEM-BN models were developed to explore the complex relationships among land use cover (crop, forest and residential land), water quality (total phosphorus, total nitrogen and dissolved oxygen), physical factors influenced by climate change (water temperature, flow velocity), other habitat characteristics (slope and substrate composition) and macroinvertebrate EPT (Ephemeroptera, Plecoptera and Trichoptera) indices (the percentage of EPT taxa, EPT richness, EPT abundance and the Shannon-Weiner Diversity Index of EPT). Three scenarios were designed to assess the possible responses of EPT indices to land use change, climate change and their interactions. Our results showed that when the change of land use and climate were considered alone, increasing crop and urban land led to declines in EPT indices whereas moderate rise of air temperature and more rainfall had opposite effects. However, the combined effect showed that the positive effects caused by climate change could weaken some negative effects, but land use change still had stronger effect on EPT indices. Our results provided more detailed understanding on how environmental stressors affect freshwater organisms, and further catchment management should integrate the combined effect of different environmental stressors on streams.

1. Introduction

Currently anthropogenic disturbances are posing a serious threat to global rivers and streams (Tejerina-Garro et al., 2005). Through land use change, anthropogenic activities have led to significant variations in inputs of nutrients, pollutants and sediments to streams, and in-stream biological and chemical processes (Elbrecht et al., 2016; Fohrer et al., 2001; Guo et al., 2016a). Rivers and streams flowing in highly agricultural landscapes more easily import more sediments, nutrients, and pesticides, which may result in greater algal production, changes in autotroph assemblage composition and declines in bank stability (Allan, 2004; Sponseller et al., 2001). Widespread urbanization and industrialization results in significant habitat degradation and biodiversity loss in rivers and streams (Roy et al., 2003; Vörösmarty et al., 2010). The increases in impervious surface area can cause erratic hydrology and large amounts of pollutants in runoff (Paul and Meyer, 2001).

Air temperature and precipitation, which are strongly associated with climate change (Karl and Trenberth, 2003; Min et al., 2011), are important factors affecting water temperature and flow velocity. Increasing air temperature has been observed from the beginning of the 20th century (Novotny and Stefan, 2007), and due to the increase of solar radiation, surface water temperatures have also experienced some rise since the 1960s in Europe, North America and Asia (Delpla et al.,

* Corresponding authors.
E-mail addresses: zhangyuan@craes.org.cn (Y. Zhang), yqwang@nankai.edu.cn (Y. Wang).

https://doi.org/10.1016/j.ecolind.2017.11.044
Received 22 December 2016; Received in revised form 20 November 2017; Accepted 21 November 2017
Available online 08 December 2017
1470-160X/

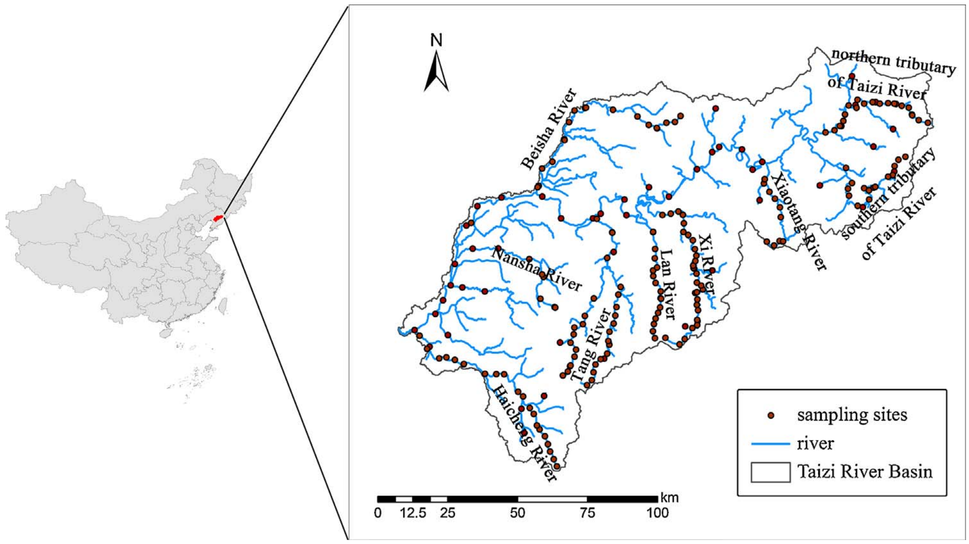

Fig. 1. Sampling sites in the Taizi River Basin, Northeastern China.

2009). Changes in precipitation are the main causes of variability in the hydrology and water balance (Novotny and Stefan, 2007). Flow velocity increases as the volume of the water in the stream increases (Ngo-Duc et al., 2007), while stream flow is generally very sensitive to precipitation with positive correlation (Groisman et al., 2001). Changes in water temperature and hydrological conditions would consequently influence water quality, habitat characteristics, and organisms living in streams (Durance and Ormerod, 2007; Guo et al., 2016b).

Stream macroinvertebrates are very sensitive to the above anthropogenic disturbances compared with other freshwater organisms, such as fish and algae (Dohet et al., 2015; Hershey and Lamberti, 1998). Their community compositions are strongly influenced by changes in stream habitats and water quality, and they have been developed as indicators to assess stream ecosystem processes and functions (Walsh, 2006). Amongst stream macroinvertebrates, Ephemeroptera, Plecoptera and Trichoptera (EPT) taxa are used for ecological studies more often compared with other taxa (Bispo et al., 2006; Peterson and Eeckhaute, 1992) because of their wide distribution, easy identification and sensitivity to disturbances. EPT taxa show lower tolerance to environmental stressors, such as land use and climate change, than other taxa. For example, EPT taxa richness is significantly lower in agricultural dominated streams than in forested streams (Lenat and Crawford, 1994). Also EPT taxa decrease as stream nutrient and contaminant levels increase in urban areas (Wilkins et al., 2015). Furthermore, macroinvertebrate abundance is predicted to reduce under warm or cool extremes (Durance and Ormerod, 2007) and the abundance and composition of EPT taxa show strong temporal variations attributed to rainfall events (Schmitt et al., 2016). Therefore, EPT taxa are often used as sensitive indicators of high-quality ecological conditions (Ferreira et al., 2014; Klemm et al., 2003).

Numerous previous studies applied traditional statistical methods, such as bivariate regression analysis (Sponseller et al., 2001) or one-way analysis of variance (Sangiorgio et al., 2014), mainly focused on the effect of a single stressor on EPT taxa without considering the complex interactions among stressors. A range of dynamic, spatially explicit and multidisciplinary models were also developed to strive for thorough quantification of ecosystem (Kragt et al., 2011), but they usually have difficulties with parameter estimation in case of limited data availability, low model transparency and excessive computation time (Landuyt et al., 2013).

Bayesian networks (BN) are increasingly used in ecological modeling and conservation in recent years (Adriaenssens et al., 2004). It is a multi-causal method with explicit constructions among diverse factors (Marcot et al., 2006), and could offer a visual framework to depict the chain of causal-effect relationships among numerous factors. BN has been used to assess the response of riparian tree species to the interactive influences of hydrological factors (Kath et al., 2016), as well as potential population response of the selected at-risk fish and wildlife species to habitat capability factors (Marcot et al., 2001). However, BN application often suffers from doubt as relies on subjective determination of structure, which is only based on expert suggestions. It would be more reliable if both expert experience and observed data could be incorporated to BNs.

Structural equation modeling (SEM) as a theoretically based method provides the assessment of relationships among multiple variables by combining causal-effect information and observed data (Grace and Pugesek, 1997; Swait, 1994). It has been used widely in social science, psychology and biology to explore the relationships among multiple variables, while the applications of SEM in ecology and environmental sciences are still limited (Grace et al., 2012; Malaeb et al., 2000). SEM is a very useful causal modeling approach, but mainly deals with linear relationships, which may result in inaccurate prediction if the relationships between independent and dependent variables are nonlinear (Gupta and Kim, 2008). Bayesian networks (BNs) could overcome these limitations. Therefore, the combination of BNs and SEM could construct the relationships between environmental factors and EPT taxa, and predict future trends under combinations of different environmental stressors.

This is the first study using the combined approaches, i.e. SEM and BNs, to assess the effect of environmental stressors on benthic macroinvertebrates. The main objectives of our study were: (1) to assess robust relationships among environmental stressors and EPT taxa by SEM, and then to estimate the interaction among these multiple types of variables by BN. (2) to predict and compare the individual and combined effects of land use and climate change on EPT taxa by SEM-BN model.

2. Materials and methods

2.1. Study area and sampling sites

The Taizi River Basin is located in Liaoning Province in Northeastern China (122°25′-124°55′E, 40°28′-41°38′N) (Fig. 1), with the drainage area 13,900 km^2. It has the temperate continental monsoon climate with long cold winters, hot rainy summers, and short springs and autumns. The mean annual precipitation is about 655–954 mm, with 50% in July and August (Kong et al., 2013; Zhang et al., 2009). The elevation in the basin increases from west to east,

with the same tendency of precipitation but a contrary trend for evaporation. The land use patterns were quite different along the stream longitudinal gradient: woodland and grass land mainly in upstream mountain and hill zones, with farmland and residential land in downstream plain land zones (Wan et al., 2013). Due to the continuous growth of population and economy, pollutants discharging into surface water keep increasing, and the Taizi River was now facing serious ecological deterioration (Bu et al., 2014).

In August 2009, 211 sites were sampled from the Taizi River Basin. Amongst those sites, 69 sites were selected by dividing the Taizi River Basin into a series of 500 km^2 grids using space gridding methods. The other 142 sites were sampled intensively from nine main tributaries of the basin, and the distance between any two nearby sites was 3 km. Those sites distributed equally along each tributary according to different lengths of tributaries (Fig. 1), 14 sites were selected from southern tributary of Taizi River, 14 from northern tributary, 10 from Xiaotang River, 13 from Lan River, 23 from Xi River, 14 from Beisha River, 29 from Tang River, 7 from Nansha River and 18 from Haicheng River.

2.2. Sample collection and processing

2.2.1. Physiochemical variables

The freshwater macroinvertebrate species were found along the river continuum and varied in sensitivity to the change of water quality (Azrina et al., 2006). Thus, ten water quality parameters were measured in our study. The water temperature (I. Temp), pH (II), electrical conductivity (III. EC), total dissolved solids (IV. TDS), and dissolved oxygen (V. DO) were all measured directly *in situ* with a handheld multiparameter water quality monitoring instrument (YSI), which was calibrated each day. At each site, two bottles of 1L water sample were collected from the monitoring reach, put into a deep freezer in the field and brought back to the laboratory within 24 h. In the laboratory, ammonia-nitrogen (VI. NH_3-N) and total nitrogen (VII. TN) (the sum of ammonia-nitrogen, nitrate-nitrogen, nitrite-nitrogen and organically bonded nitrogen) were measured using Nessler's reagent and the alkaline potassium persulfate oxidation-UV spectrophotometric method, respectively. Total phosphorus (VIII. TP) was measured after phosphates were converted to orthophosphate using concentrated nitric and sulfuric acid with the ammonium molybdenum blue method (Murphy and Riley, 1962). The permanganate index (IX. COD_{Mn}) was measured using the potassium permanganate method (Dixon et al., 2011) and 5-day biological oxygen demand (X. BOD_5) was determined by quantifying the dissolved oxygen before and after the 5-day incubation at 20 °C. Sampling, preservation, transportation and analysis of all water samples were performed following national quality standards for surface waters, China (GB3838-2002) (China, 2002).

The latitude, longitude and altitude of each site were measured using a handheld global positioning system (GPS). In addition, flow velocity and depth measurements were made using a current meter at each site. The slope of the study area was calculated using the slope tool in the ArcGIS system based on the elevation dataset.

The substrate composition was an important factor affecting macroinvertebrate distribution. Substrate measurements were processed at each site and weighted substrate percentages were summed to transform into a single substrate index (Jowett and Richardson, 1990). The weighting values were based on the original Instream Flow Incremental Methodology (IFIM) substrate codes and the research of Ian G. Jowett (Jowett et al., 1991) as follows:

$$\text{Substrate index} = 0.08\%\text{bedrock} + 0.07\%\text{boulder} + 0.06\%\text{cobble} + 0.05\%\text{gravel} + 0.04\%\text{finegravel} + 0.03\%\text{sand}$$

2.2.2. Benthic invertebrate assemblages

Numerous techniques were available for quantitative sampling of organisms living in the sediment or gravel of stream beds, in which the Surber sampler was recommended as the easiest method (Rempel et al., 2000). In this study, macroinvertebrates were qualitatively and quantitatively sampled in each site using a Surber sampler (30 cm × 30 cm, 1 mm mesh size). According to previous studies (Boothroyd and Stark, 2000; Watanabe et al., 2014), the number of replicates between three and six should be adequate to characterize most macroinvertebrate communities. Three replicates were collected randomly at each site in this study, which in a manner represented the longitudinal and cross-sectional heterogeneity of the riffle. All samples were preserved in 70% ethanol in the field, and were sorted, counted, and identified to species or genus in the laboratory (Elliott et al., 1988; Merritt and Cummins, 1996; Wiggins, 2015).

A total of 157,502 macroinvertebrates were collected, which were classified into 24 orders, 69 families, 101 genera and 123 species. The richness of Ephemeroptera, Plecoptera and Trichoptera were 32, 11 and 37, and they accounted for 14.47%, 0.59% and 12.09% of the total abundance, respectively.

Four EPT indices were used as the primary metrics to assess compositional differences of macroinvertebrate communities among sampling sites, i.e., the percentage of EPT taxa ('percent'), EPT richness ('richness'), EPT abundance ('abundance') and the Shannon-Weiner Diversity Index of EPT ('H'). The percentage of EPT taxa can show the proportion of macroinvertebrates intolerant to poor water quality, and it is often used to assess the water quality in biological monitoring (Ab Hamid and Rawi, 2014; Narangarvuu et al., 2015). EPT richness and EPT abundance played important roles in assessing stream biodiversity (Konrad et al., 2008; Pond, 2012). The Shannon-Weiner Diversity Index of EPT was one of the mathematical measures of species diversity in a community, which took into account both species richness and species evenness under specific assumptions (Jhingran et al., 1989).

2.2.3. Land use

Land use data were interpreted from SPOT 5 images before crops were harvested in September 2010. The images pre-processing, such as image correction and connection, were processed by ERDAS (Zhang et al., 2013). The classification of land use type was mainly determined by visual interpretation, and field survey was taken for some uncertain types. The study area had five dominant land use types, including woodland, farmland, residential land, grass land and barren land, which covered 57.20%, 30.11%, 8.56%, 3.24% and 0.48% of the total drainage area, respectively (Fig. 2). The surface area contributing to drainage through each sampling point was quantified by a digital elevation map based on the geographical information system (GIS), and the area of each land use type were then identified for each sampling reach.

3. Model development

In this study, we developed a conceptual model to explore the effect of land use and climate change on macroinvertebrate EPT by SEM, then built BNs on the same conceptual model to predict and diagnose the changes in macroinvertebrate EPT induced by varying land use and climate change.

3.1. Structural equation models

A simple conceptual cause-effect model was constructed at first (Fig. 3). Four latent variables were considered, i.e., 'land use', 'physical' and 'water quality' represented environmental stressors, and 'EPT indices' indicated macroinvertebrate communities. The land use factor included the percent of farmland, woodland, residential land, grass land and barren land for each sampling site. Physical factors indicated the variables associated with macroinvertebrates habitat, including substrate index, slope, depth, velocity, altitude, water temperature, pH, longitude and latitude. We selected TN, TP, DO, and COD_{Mn} as

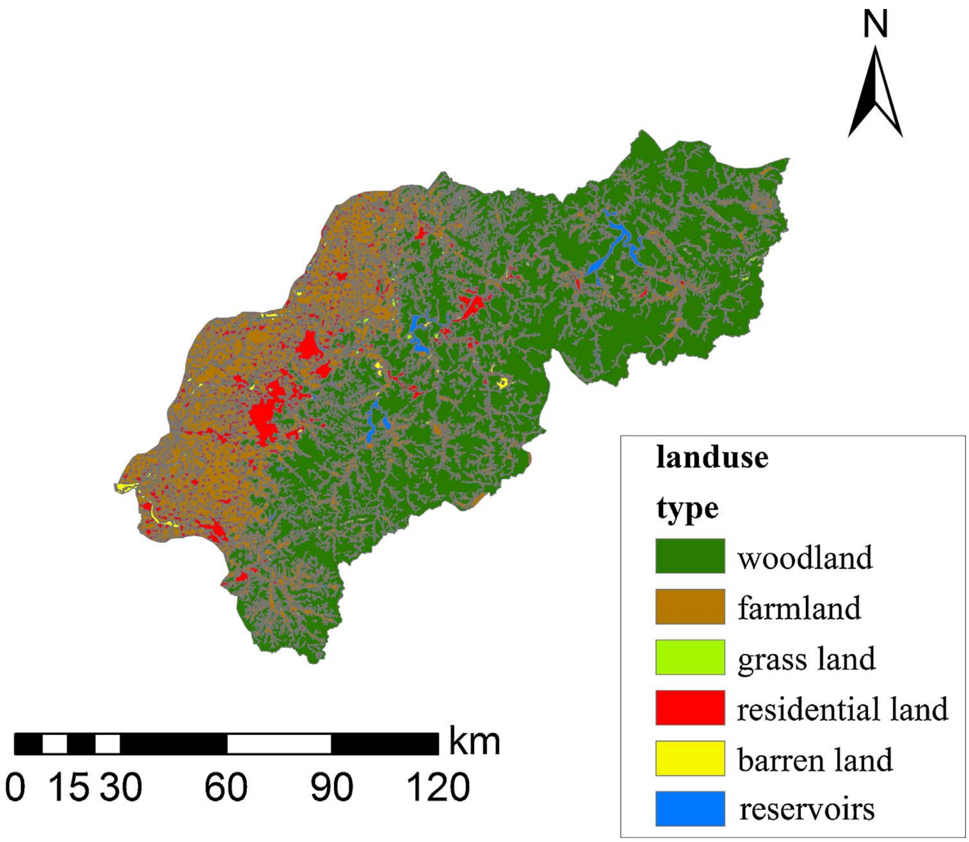

Fig. 2. Land use types in Taizi River Basin, Northeastern China.

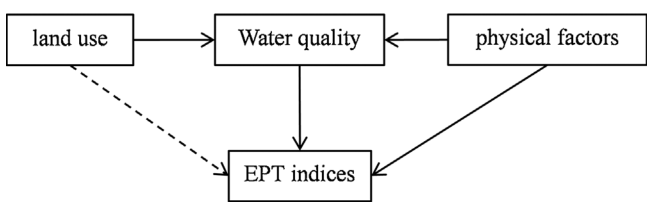

Fig. 3. The conceptual cause-effect relationships between latent variables.

preliminary explanatory variables to represent 'water quality' in the river. Four EPT indices ('percent', 'richness', 'abundance' and 'H') were represented as 'EPT indices'.

The conceptual model illustrated a hypothesis that EPT indices were directly influenced by three factors, and water quality was controlled by land use and physical factors. The primary structure was corrected repeatedly in the training procedure. Each latent variable was linked with several observed variables. The Pearson correlation analysis was applied on observed data of each latent factor and the significantly correlated variables were eliminated.

Four measurement models were constructed for latent variables to select the most fitted observed variables. The SEM models were fit using Amos (Amos Version 21.0, IBM/SPSS, Chicago, Illinois, 2012) with standardized data. The variables with different orders of magnitude may lead to inaccurate results and pretreatment was applied to eliminate the effect of orders. Based on the results of initial models, some modifications were performed to lead a better fit. For example, COD_{Mn} was eliminated from the water quality model, and EPT abundance was not included in the final 'EPT indices' model. Table 1 lists the summary statistics of variables finally used in the SEM model.

After all measurement models were satisfied with χ^2 statistic and root mean square error of approximation, the conceptual model and the measurement parts were combined into one structural equation model. Maximum likelihood method was used to estimate the coefficients of paths between each pair of variables. Results showed that the correlation between 'land use' and 'EPT indices' was not significant, so we removed the path between them (the dotted line in Fig. 3). The pathways between crop and TP, temperature and DO were added into the final model according to the modified index (MI) and previous research (Fig. 4). When the structural equation model could satisfy a series of statistical criteria, it indicated that the structural basis of the model was reliable and further research could be carried out based on it.

3.2. The bayesian network

After the cause-effect relationships between environmental stressors and EPT indices were identified by SEM, BN was established for model prediction and diagnosis. The two most important steps to develop a BN were structure learning and conditional probability estimation. The structures of the BNs are usually generated from experiences, which may lead to spurious relationships. In this study, NETICA (Norsys Software Corp.) software was used to develop the BNs based on the same structure exported from the SEM. Prior to parameter learning, all variables were needed to be discretized into a relevant number of classes to specify the conditional probability tables (CPTs) of the nodes. The self-organizing map (SOM) method was used to avoid subjective effect in clustering the variables. The SOM as a pattern recognition technique incorporates unsupervised learning and a type of neural network (Kohonen, 1998). In this study, EPT indices and all environmental variables except for TN and TP were classified were classified into four categories using the SOM Toolbox in MATLAB. TN and TP concentrations were divided based on the Chinese system of Environmental Quality Standards for Surface Water [The National Standards of the People's Republic of China: Environmental Quality Standard for Surface Water (GB 3838-2002), 2002]. To evaluate the predictive power of the BN, a part of the data set (80%) was used as training data to develop the BN and the remaining (20%) was used as validating data to evaluate the model (Witten and Frank, 2005). The range and sample

Table 1
Summary statistics for variables used in structural equation models.

latent variable	observed variable	definition	mean	SD	min	max
landuse	forest (%)	% woodland + % grass land in the basin	55.74	29.19	0	94.53
	building (%)	% residential land in the basin	5.91	8.70	0	69.85
	crop (%)	% farmland in the basin	33.76	20.88	0	89.15
physical	temperature	water temperature (°C)	19.25	4.63	6.80	28.20
	slope	the angle to the horizontal (degree)	7.27	8.21	0	50.08
	velocity	the flow volume in unit time (m/s)	0.36	0.20	0	0.83
	substrate	substrate index	1.81	1.22	0.004	5.35
water quality	TP	the total phosphorous concentration (mg/l)	0.15	0.27	0	2.63
	TN	the total nitrogen concentration (mg/l)	4.90	4.52	0.44	22.60
	DO	the dissolved oxygen concentration (mg/l)	10.13	4.42	0.01	25
EPT indices	richness	EPT richness	9.45	8.35	0	29
	H'	the Shannon-Wiener Diversity Index of EPT	1.86	1.22	0	4.53
	percent	the percent of EPT taxa (%)	28.69	27.58	0	92.24

number of each state for variables are shown in Table 2. The four latent variables were divided into three states as uniform distribution. Once the conceptual model was developed and data were prepared, the BN was calibrated and used for inference.

Several algorithms with different complexity, computational efficiency and generality were considered for BN inference. In this study, the expectation–maximization (EM) algorithm was chosen. It is an iterative method to find maximum likelihood and could be applied on the model with unobserved latent variables (Neal and Hinton, 1998). Our model had four latent variables with no observed data and was suitable to use EM algorithm.

3.3. Scenario analysis

When the BN model was validated, the relationships between parent variable and child variable were represented as probabilities in the CPTs. The parentless nodes were described by marginal probability distributions, while the conditional probabilities of child nodes were allocated for each combination of states in their parent nodes (Allan et al., 2012). For an example, Table 3 provided the CPT of TP with only one parent node (crop). The sites with low percent of crop (state I and II) were likely to have lower TP concentration, whereas it was more uncertain to predict the TP concentration when the crop percent increased. By changing the CPTs of input variables, the relationships between EPT indices and multiple drivers related to land use and physical factors could be predicted. For example, when 'crop' was assumed to increase and set to state IV, the CPT of TP was 0.1, 0.4, 0.4 and 0.1 according to Table 3.

As two main factors influencing the macroinvertebrate communities, human activities were often directly reflected by the change of land use (Meyer and Turner, 1992), while changes in the climatic conditions could be considered as a consequence of the effect came from natural processes (Parmesan and Yohe, 2003). We set three scenarios to explore the effects of land use change, climate change and the combined effects on EPT indices, respectively. Taizi River Basin was located in a fast-developing region under tremendous pressure from human activities. The area of residential land was predicted to increase based on historical remote sensing images (Liu et al., 2011). The amount of crop land per capita in China was much less than the world average, and would continue to decline with increasing population (Khan et al., 2009). Facing the challenge of food production, we assumed that the area of cropland would increase in the future. So Scenario 1 was defined by only altering the node states of land use and assumed that the area of crop and urban land would increase, whereas the forestland would decrease in the future, according to the current development trends in the study basin. Besides land use, climate change also showed significant effects on macroinvertebrate EPT communities. Gao et al. (Gao et al., 2011) studied the regional climate change of Northeast China and predicted that the air temperature would be higher and precipitation would increase in the future. Thus the water temperature was assumed to rise, and the velocity would be accelerated in scenario 2. Scenario 3 considered changes in both land use and climate, and explored the combined effect. The scenario outcomes were evaluated by running the BN model and found out the changes in node probabilities of EPT indices.

4. Results

4.1. Model performance

The final structural equation model provided a good fit to the data. The χ^2-test statistic value was 130.40 with 58 df (degrees of freedom) and non-significant *p*-value 0.22, indicating that the model was plausible. The root mean square error of approximation (RMSEA) was 0.07, while Browne and Cudeck argued that a value of about 0.08 or less for the RMSEA suggested a close fit based on their experience (Browne and Cudeck, 1993). The goodness of fit index (GFI) and adjusted goodness of fit index (AGFI) range between 0 and 1 and a value close to 1 indicates a good fit. The values of the GFI and AGFI in this stufy were 0.92 and 0.87, respectively. These two indices was both measures of fit between the hypothesized model and the observed covariance matrix (Tanaka and Huba, 1985), and AGFI is affected by the number of indicators of each latent variable (Baumgartner and Homburg, 1996). Further, all path coefficients calculated from the model were significantly different from zero ($p < 0.05$).

The predictive capacity of the BN was evaluated by the correct rates (CR) of the output variables (i.e. richness, *H'*, and percent) during testing period (Table 4). Results indicated that the correct classification rates of 'richness' and 'H' were higher than 'percent', mainly showing a high trend towards uncertain classifications for 'percent' class prediction.

The BN predictions for the current situation are shown in Fig. 5. The initial model represented the starting point for exploring the potential effects of multiple factors on EPT indices. In order to rank the importance of variables, the sensitivity analysis (SA) of the three output nodes ('richness', 'H' and 'percent') to other nodes was investigated (Table 5). It can be used to determine the relative influences from other nodes to specified nodes. Sensitivity in the BN is defined as the expected reduction in variation (VR) of some specific variables. The key pollution indicator was TN in the study area as 68.7% of the sites had TN concentration higher than 2.0 mg/l, but TP had the strongest impact on all three output nodes. The nodes 'crop', 'DO' and 'TN' had stronger effects than the other variables. The result presumably reflected the direct relationship between water quality and EPT indices. The nodes 'slope' and 'bottom' had the lowest importance to the output nodes (Table 5) indicated that regional scale EPT indices were relatively insensitive to topography and substrate. The impacts of other land use and physical factors were shown as 'temperature' > 'building' > 'forest' > '

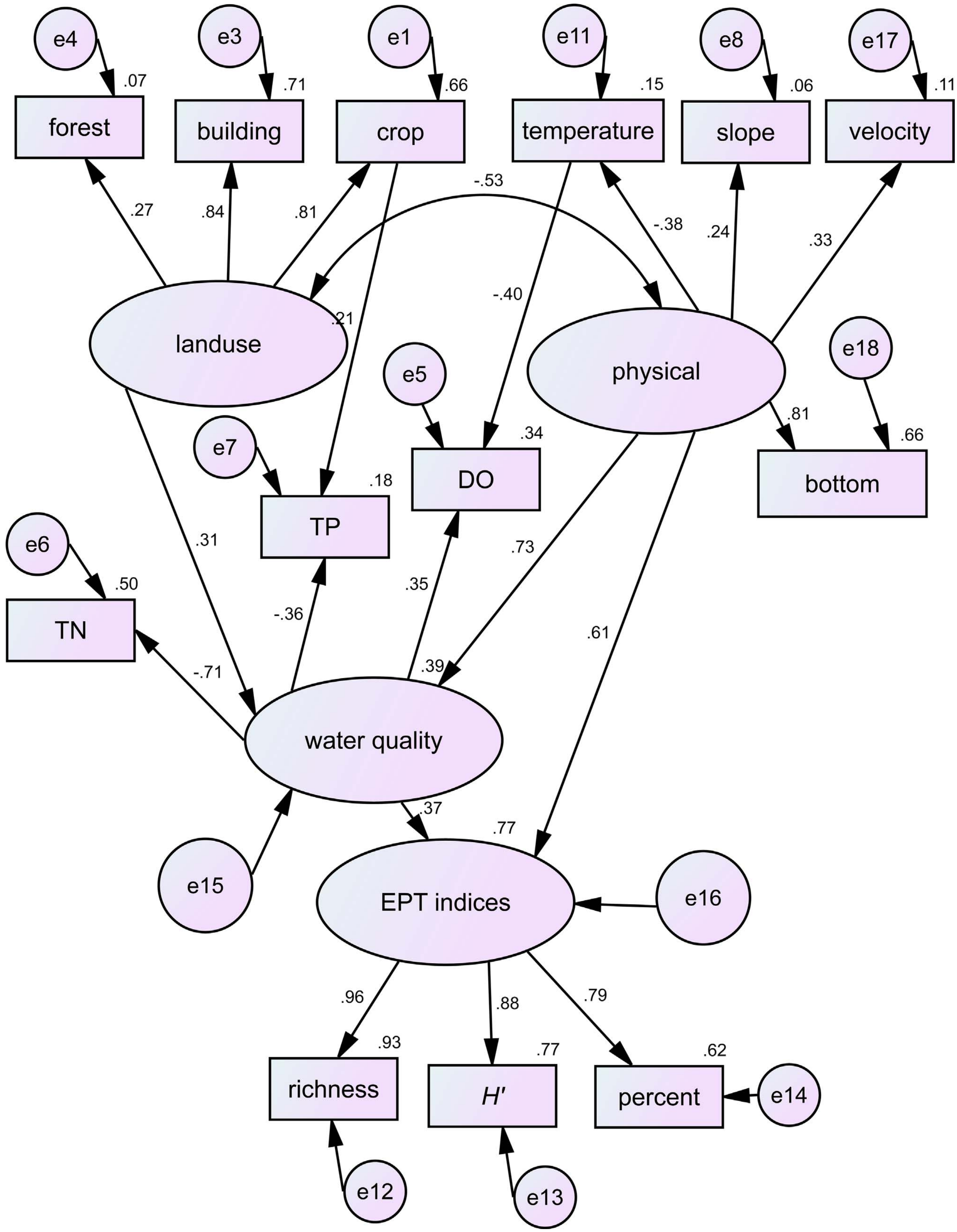

Fig. 4. Final structural equation model (SEM) for Taizi River Basin.

Table 2
The range and sample number of each state for variables used in Bayesian Network during training period.

Variables	I		II		III		IV	
	range	number	range	number	range	number	range	number
crop	0–37.70	117	39.99–66.14	36	66.19–79.05	10	79.86–89.15	8
forest	0–59.92	91	60.87–79.48	48	79.88–85.18	23	85.46–94.53	9
building	0–6.10	120	6.37–15.23	34	15.34–18.86	8	19.83–69.85	9
temperature	6.8–15.6	37	15.7–17.9	22	18.0–20.8	39	21.0–28.2	73
slope	0–3.45	66	3.60–7.74	49	8.30–14.55	32	15.01–50.08	24
velocity	0.01–0.21	46	0.22–0.35	37	0.36–0.48	43	0.49–0.83	45
substrate	0.004–1.58	74	1.60–2.07	37	2.10–2.77	25	2.85–5.35	35
TP	0.001–0.09	106	0.10–0.19	34	0.20–0.37	16	0.40–2.63	15
TN	0.44–0.99	6	1.01–1.47	23	1.55–1.99	25	2.00–22.60	117
DO	0.01–8.00	59	8.30–10.69	39	10.76–13.50	59	13.78 – 25.00	14
richness	0–4	55	5–9	39	10–16	38	17–29	39
H'	0–0.01	18	0.02–0.99	24	1.00–1.79	31	1.81–4.53	98
percent	0–11.54	64	12.28–30.59	35	33.57–54.02	29	54.51–92.24	43

Table 3
An example of the conditional probability table of TP with one parent node (crop).

crop	TP			
	I	II	III	IV
I	0.72	0.13	0.07	0.08
II	0.67	0.33	8.33E-07	8.33E-07
III	0.40	0.31	0.13	0.16
IV	0.10	0.40	0.40	0.10

Table 4
The correct rate of classification for validation data.

variables	correct	error	correct rate
richness	36	4	90%
H'	33	7	82.5%
percent	31	9	78.5%
Total	100	20	83.3%

velocity'. The sensitive ranking results of variables impacted on water quality were consistent with the ranking sequence of variables on EPT indices. The node 'crop' showed the strongest impact, reflecting the significant correlations among agriculture, nutrients and aquatic

Table 5
The sensitivity analysis results (variation reduction values) of three output nodes to other nodes.

Nodes	richness	*H'*	percent
TP	0.086	0.068	0.062
crop	0.026	0.017	0.018
DO	0.013	0.0075	0.0098
TN	0.011	0.0073	0.0083
temperature	0.0031	0.0017	0.0022
building	0.0027	0.0016	0.0019
forest	0.0013	0.00024	0.00095
velocity	0.00061	0.00013	0.00073
slope	0.0002	0.00017	0.00011
bottom	6.97E-05	7.28E-06	8.17E-05

insects. The nodes 'temperature', 'building', 'forest' and 'velocity' had moderate effects, while 'slope' and 'bottom' had weak influences. These findings indicated a stronger influence of land use than physical factors on both water quality and EPT indices, which would be helpful for the design and analysis of predictive scenarios.

4.2. Scenario analysis results

A BN was used by entering 'evidence' (i.e., changes in node

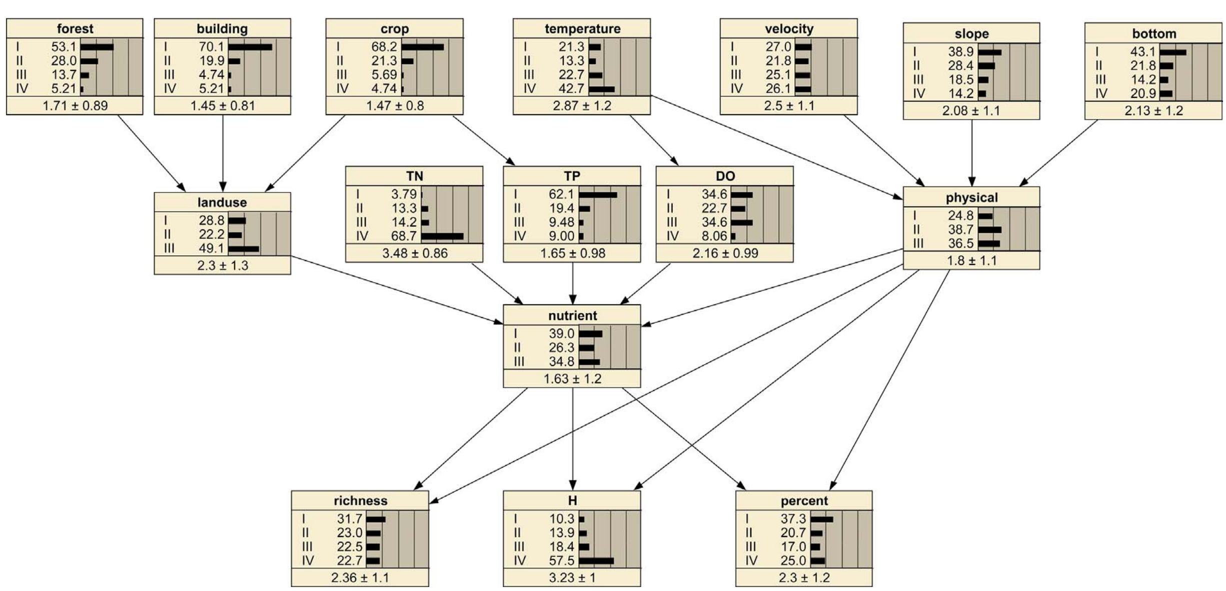

Fig. 5. The Bayesian Network (BN) for Taizi River Basin.

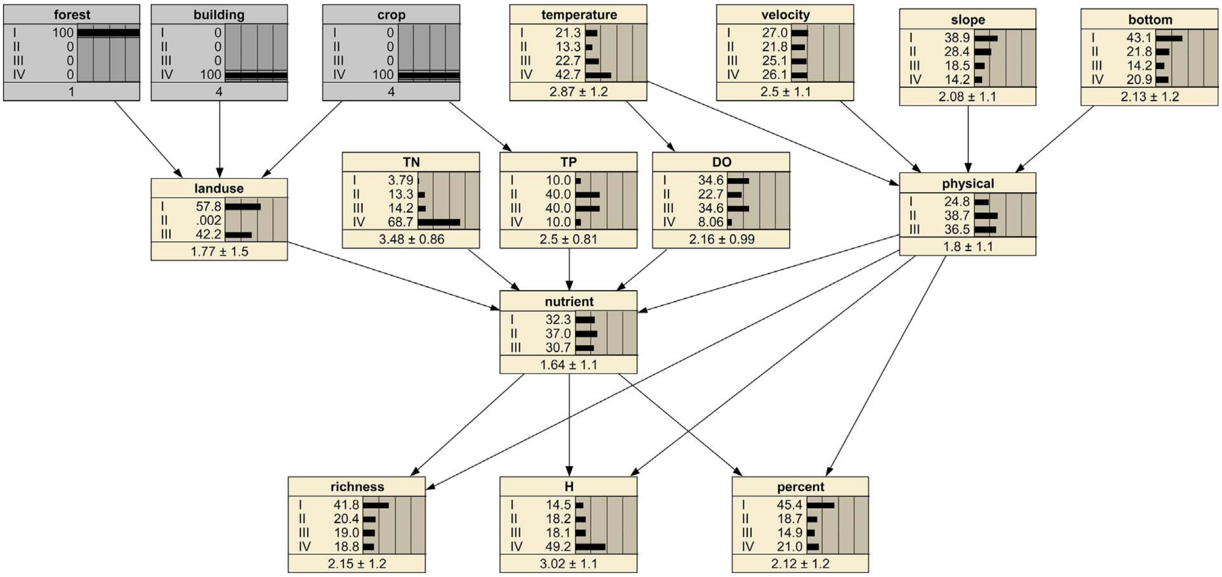

Fig. 6. Scenario 1: the effect of land use change on EPT indices.

probabilities) and generating different scenarios to show the effect of interactions between nodes. Among the three latent variables that would affect the EPT indices, 'physical factor' and 'nutrient' were the most significant ones. The node 'land use' was not directly linked to EPT indices, but the change of land use could reflect anthropogenic activities and influence water quality.

The first scenario mainly reflected the effect of land use change on EPT indices. Under this scenario, the farmland and residential land were assumed to be expanding, and forest land may decrease in the future (forest = state I, crop = state IV, building = state IV). As the percent of the specific land use increased from state I to state IV, the BN for scenario 1 was set as Fig. 6. The CPTs of other nodes remained the same with the initial model, so the change of output states could be analyzed. The state I to IV for 'richness', 'H' and 'percent' indicated the trend of ecological status from bad to good (Table 2). Under scenario 1, the percent of state I and II for 'richness' increased 5.6% and 0.5%, respectively. The percent of node 'H' increased 2.1% and 2.3% for states I and II, respectively. The state I of 'percent' increased 5.1% indicated the decrease of EPT taxa among macroinvertebrate communities.

The second scenario was aimed to find out the effect of climate change on EPT indices. We assumed that the air temperature would be higher and precipitation would increase in the future. Under this assumption, the change of climate may cause the rise of water temperature and the acceleration of flow rate (temperature = state IV, velocity = state IV). The BN for scenario 2 was set as Fig. 7. Under this scenario, the percent of state I for 'richness', 'H' and 'percent' decreased 6.2%, 2.6% and 4.9%, while the percent of state IV increased 2.4%, 4.9% and 2.4%, respectively. Both the species richness and the amount of EPT taxa would increase under scenario 2, which indicated a suitable habitat under the warming and rainy conditions.

The third scenario was set to study the combined effects of land use change and climate change. Under this scenario, 'crop', 'building', 'temperature' and 'velocity' were thought to be increased, while 'forest' was assumed to decrease (forest = state I, crop = state IV, building = state IV, temperature = state IV, velocity = state IV) (Fig. 8). The results were similar to scenario 1 but with less variation. The three EPT indices all trended worse under scenario 3. The percent of state I for 'richness', 'H' and 'percent' increased 1.0%, 0.3% and 1.2%, while

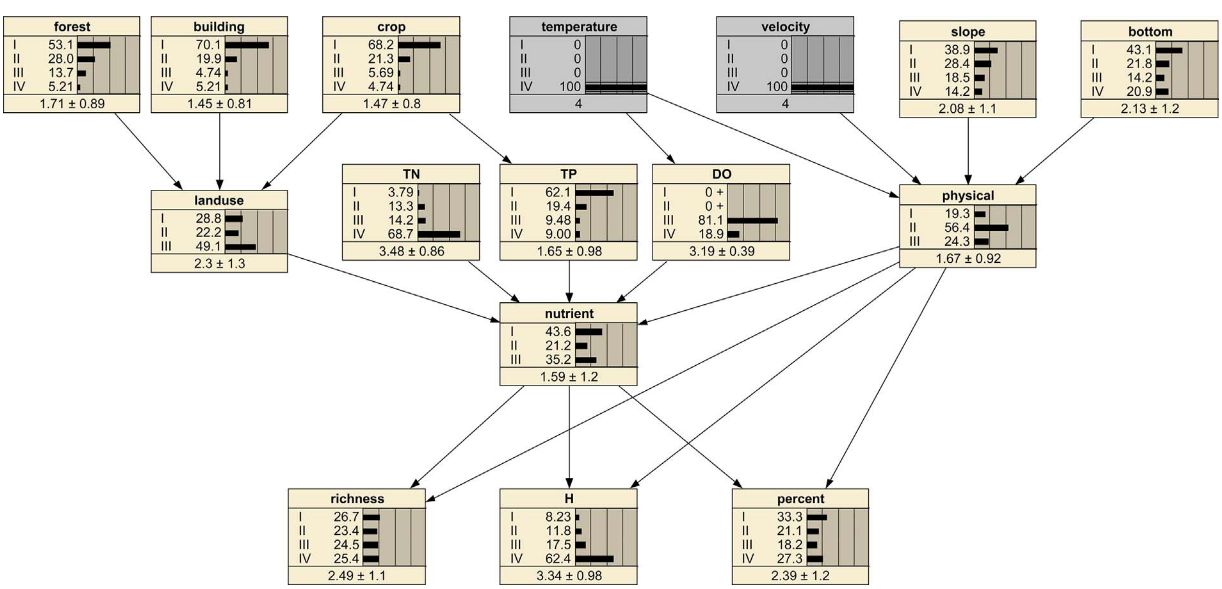

Fig. 7. Scenario 2: the effect of climate-change on EPT indices.

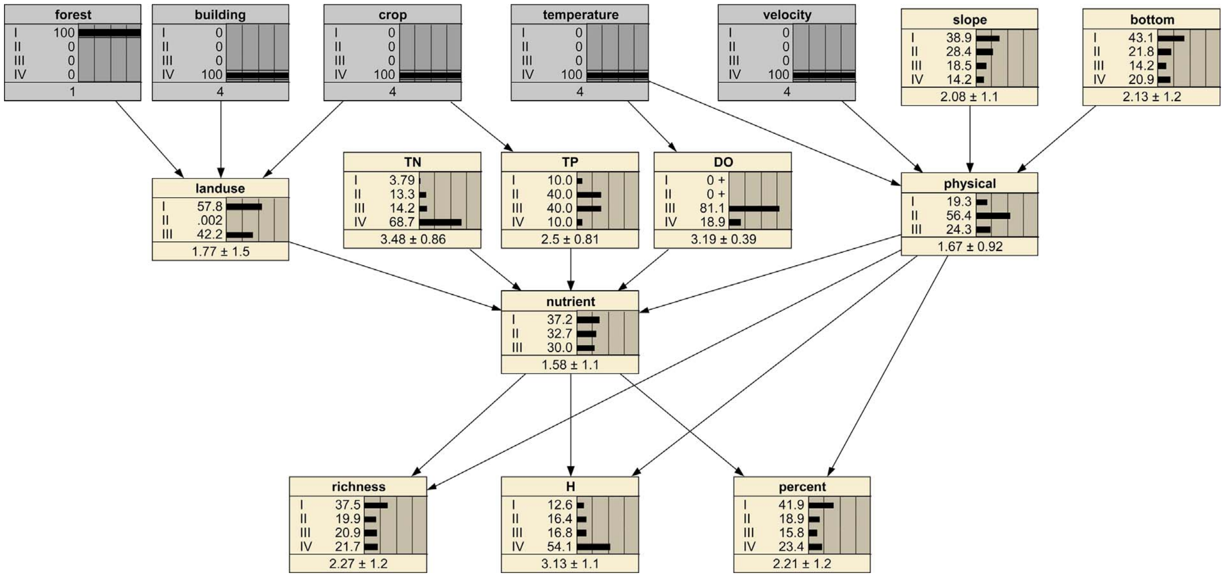

Fig. 8. Scenario 3: the combined effects of land use and climate change on EPT indices.

the percent of state IV decreased 1.4%, 1.5% and 1.1%, respectively. The results indicated a degraded tendency of water quality and stream habitat.

The variation in responses among the three scenarios are shown in Fig. 9. The results of scenario 3 showed a detrimental effect on 'EPT index', as richness, percent and Shannon diversity decreased, which showed a same tendency of scenario 1. However, the lesser degree of deterioration in scenario 3 indicated that the effects of changes in land use and climate were contrasting so that they just cancelled each other out to some extent and their combined effect was not strong.

5. Discussion

Numerous studies have used either SEM or BNs to investigate relationships among ecological variables, or to evaluate potential effects of alternative management decisions (Chen and Pollino, 2012; Malaeb et al., 2000; Marcot et al., 2001; McCann et al., 2006). However, the actual and suspected links included in the cause-effect diagram of BNs were easy to be confused. The structure of the model in this study was designed in two steps by linking SEM to BNs. The conceptual model exported from the SEM, which had satisfied a series of statistical criteria (such as the values of χ^2, GFI, AGFI and RMSEA), were then calibrated and validated by the BN. In comparison with elicitation approach depending on the discussion of modelers and ecologists, the integrated method took more account of the observed data and was more oriented toward the study area. The predictors chosen by this method were relatively low in number, which could avoid the over-fitting problem of the BN, especially when the dataset was small (Cheng and Greiner, 1999). The excluded variables were tested by the SEM and considered as the least sensitive ones to describe the taxon richness and

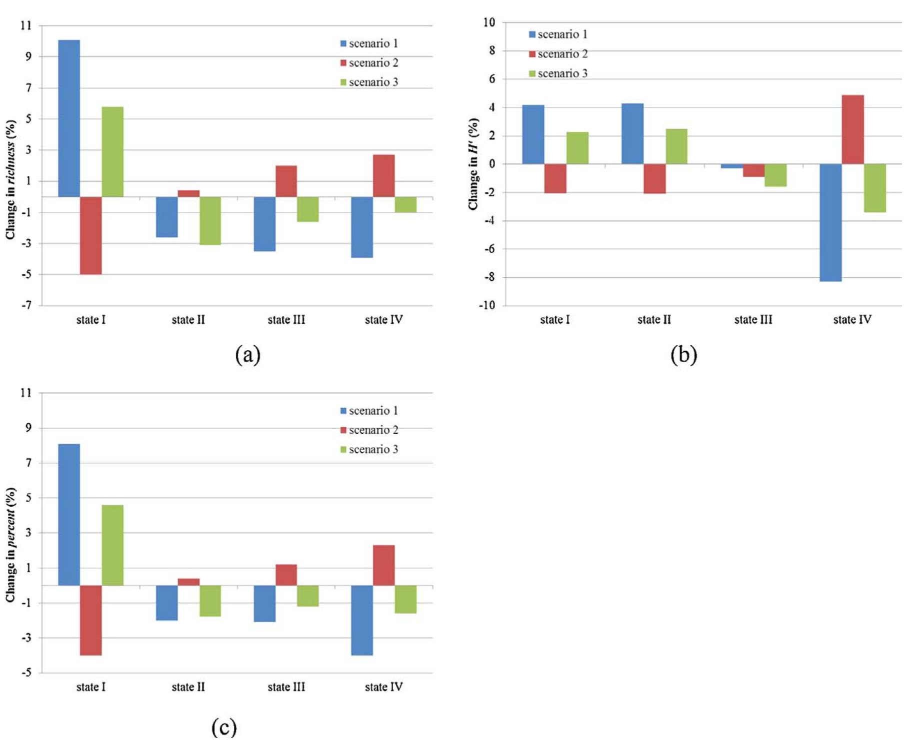

Fig. 9. Changes of three scenarios in (a) richness, (b) *H*', (c) percent for each state.

distributions. The prediction estimated by the model with data-based construction would be beneficial for the exploration of the effects of land use and climate change on the macroinvertebrate community in the study basin.

In our study, the scenario 1 aimed to explore the effects of land use change in the study basin. Previous studies had demonstrated that forestland had a positive contribution to water quality, while cropland and residential land had negative contribution to water quality (Lenat and Crawford, 1994; Ometo et al., 2000; Tong and Chen, 2002). The results of scenario 1 were consistent with such previous findings. The increased percent of state I and II for 'richness' indicated a decrease of EPT species types, which usually implied the disruption to stable ecosystems (Dewalt et al., 1999). The change of Shannon–Wiener index under scenario 1 implied the decrease of both EPT species types and individuals (Bispo and Oliveira, 2007). The change of 'percent' shown in Fig. 6 implied more severe polluted rivers. The results of all the three output variables indicated that the assumed change of land use would lead to the deterioration of living habitat and decrease of both EPT richness and diversity.

All of the three EPT indices tended to be better under scenario 2. The mean water temperature of Taizi River in August was 19 °C and the state IV was between 21 °C to 28 °C. The surface water was relatively cold in Northeast China, while most benthic taxa were tended to track suitable habitats under the warmer conditions (Kuemmerlen et al., 2015). The moderate rise of water temperatures in Taizi River would be beneficial to the macroinvertebrate communities, as previous studies indicated a strong positive relationship between feeding rates and metabolism with temperature variation for insect communities (Lessard and Hayes, 2003). The accelerated flow rate was good for the transfer of nutrients and other pollutants (Li et al., 2017), which would help to maintain stable water quality and habitat. The assumed change in climate conditions indicated a better living habitat for macroinvertebrate communities in Taizi River Basin.

The scenario 3 considered the combined influence of land-use change along with climate change on EPT indices. Climate and land use often interact in ways, which in turn influence biodiversity (Hansen et al., 2001). The result of scenario 3 also showed that the joint influence of climate and land use was not a simple linear relationship. The positive effects caused by climate change could weaken some negative effects, but land use change still had a much greater effect on the distribution and amount of EPT taxa. Even if the change of land use took place in a warm and rainy year, it would lead to an adverse change of macroinvertebrate communities in Taizi River Basin.

The integrated method used in this study aimed to decrease the subjective biases of expert elicitation of structure and probabilities in a BN model (Chen and Pollino, 2012). However, uncertainties associated with the input nodes and parameters still existed and were propagated through the network to the final model endpoints. Due to the limitation of samples, the number of sites clustered in some states (such as state IV of crop, state III of building, etc.) was too small to calibrate the BN model, which caused the uncertainties reflected in the CPTs of EPT variables. Though the representation of uncertainty could help decision makers to identify risks of undesirable outcomes associated with management alternatives (Cain, 2001), the update with newly data collected from future monitoring programs would improve the network.

6. Conclusion

In this study, we established a model that linked SEM and BN to explore the effects of environmental factors (i.e., land use and climate change) on macroinvertebrate EPT indices. The SEM model was calibrated using the data set of Taizi River Basin from 211 sites to provide a valid assessment of related structure among the studied variables. Then a BN was established based on the same structure and data to support prediction and decision-making. The model accuracy and potential errors in the network were evaluated and identified by CPTs for each node and sensitivity analysis.

Three scenarios were implemented to estimate the effects of land use change, climate change and the combined effects on EPT indices, respectively. The results showed that land use change under increasing anthropogenic pressures had negative effects on the stream macroinvertebrate communities, while the climate change would have some positive effects in the near future in Taizi River Basin. The combined scenario with a less detrimental effect on EPT indices than land use change indicated a stronger effect of land use change than climate change. The expansion of crop and urban land would lead to decreases in diversities and EPT taxa even in a warm and rainy year.

The BN models could not reflect exact ecosystem processes, because even the most accurate BN is still just a model of the ecosystem structure. However, the transparent model could improve the understanding of cause-effect relationships between macroinvertebrates and the main influencing environmental stressors. Our results indicated that the outcome of BN would be improved by adding newly collected data from future monitoring programs, given the existing model uncertainties. The three scenarios set based on our model will be helpful to assess effects of future management strategies on ecosystems before they are carried out.

Acknowledgements

This research is funded by the National Natural Science Foundation of China (Grant No. 41571050) and the Technology Research and Development Program of the Higher Education Institutions of Tianjin (043135202JW1716). Comments and suggestions from five anonymous reviewers and the editor are greatly appreciated.

References

Ab Hamid, S., Rawi, C.S.M., 2014. Ecology of Ephemeroptera, Plecoptera and Trichoptera (Insecta) in rivers of the Gunung Jerai Forest Reserve: diversity and distribution of functional feeding groups. Trop. Life Sci. Res. 25, 61–73.
Adriaenssens, V., Goethals, P., Charles, J., De Pauw, N., 2004. Application of Bayesian Belief Networks for the prediction of macroinvertebrate taxa in rivers. Ann. de Limnol.-Int. J. Limnol. 181–191 (Cambridge Univ Press).
Allan, J.D., Yuan, L.L., Black, P., Stockton, T., Davies, P.E., Magierowski, R.H., Read, S.M., 2012. Investigating the relationships between environmental stressors and stream condition using Bayesian belief networks. Freshw. Biol. 57, 58–73.
Allan, J.D., 2004. Landscapes and riverscapes: the influence of land use on stream ecosystems. Annu. Rev. Ecol. Evol. Syst. 35, 257–284.
Azrina, M., Yap, C., Ismail, A.R., Ismail, A., Tan, S., 2006. Anthropogenic impacts on the distribution and biodiversity of benthic macroinvertebrates and water quality of the Langat River, Peninsular Malaysia. Ecotoxicol. Environ. Saf. 64, 337–347.
Baumgartner, H., Homburg, C., 1996. Applications of structural equation modeling in marketing and consumer research: a review. Int. J. Res. Market. 13, 139–161.
Bispo, P.C., Oliveira, L.G., 2007. Diversity and structure of Ephemeroptera, Plecoptera and Trichoptera (Insecta) assemblages from riffles in mountain streams of Central Brazil. Revista Brasileira de Zoologia 24, 283–293.
Bispo, P.C., Oliveira, L.G., Bini, L.M., Sousa, K.G., 2006. Ephemeroptera, Plecoptera and Trichoptera assemblages from riffles in mountain streams of Central Brazil: environmental factors influencing the distribution and abundance of immatures. Braz. J. Biol. 66, 611–622.
Boothroyd, I., Stark, J., 2000. Use of invertebrates in monitoring. New Zealand stream invertebrates: ecology and implications for management. pp. 344–373.
Browne, M.W., Cudeck, R., 1993. Alternative ways of assessing model fit. SAGE Focus Ed. 154 (136-136).
Bu, H., Meng, W., Zhang, Y., Wan, J., 2014. Relationships between land use patterns and water quality in the Taizi River basin, China. Ecol. Indic. 41, 187–197.
Cain, J., 2001. Planning improvements in natural resource management. Guidelines for Using Bayesian Networks to Support the Planning and Management of Development Programmes in the Water Sector and Beyond. Centre for Ecology and Hydrology.
Chen, S.H., Pollino, C.A., 2012. Good practice in Bayesian network modelling. Environ. Modell. Softw. 37, 134–145.
Cheng, J., Greiner, R., 1999. Comparing bayesian network classifiers. In: Proceedings of the Fifteenth Conference on Uncertainty in Artificial Intelligence. Morgan Kaufmann Publishers Inc. pp. 101–108.
China, E., 2002. Environmental Quality Standards for Surface Water. MEP, China, pp. GB3838–2002.
Delpla, I., Jung, A.-V., Baures, E., Clement, M., Thomas, O., 2009. Impacts of climate change on surface water quality in relation to drinking water production. Environ. Int. 35, 1225–1233.
Dewalt, R., Webb, D., Harris, M., 1999. Summer Ephemeroptera, Plecoptera, and Trichoptera (EPT) species richness and community structure in the lower Illinois

River basin of Illinois. Great Lakes Entomol. 32, 115–132.
Dixon, J., Wallace, A.M., O'Toole, S., Ahmed, S.F., 2011. Water quality. determination of permanganate index (ISO 8467:1993). J. Avian Biol. 42, 505–513.
Dohet, A., Hlúbiková, D., Wetzel, C.E., L'Hoste, L., Iffly, J.F., Hoffmann, L., Ector, L., 2015. Influence of thermal regime and land use on benthic invertebrate communities inhabiting headwater streams exposed to contrasted shading. Sci. Total Environ. 505, 1112–1126.
Durance, I., Ormerod, S.J., 2007. Climate change effects on upland stream macro-invertebrates over a 25-year period. Glob. Change Biol. 13, 942–957.
Elbrecht, V., Beermann, A.J., Goessler, G., Neumann, J., Tollrian, R., Wagner, R., Wlecklik, A., Piggott, J.J., Matthaei, C.D., Leese, F., 2016. Multiple-stressor effects on stream invertebrates: a mesocosm experiment manipulating nutrients, fine sediment and flow velocity. Freshw. Biol. 61, 362–375. http://dx.doi.org/10.1111/fwb.12713.
Elliott, J.M., Humpesch, U.H., Macan, T.T., 1988. Larvae of the British Ephemeroptera: a Key with Ecological Notes. Freshwater Biological Association.
Ferreira, W.R., Ligeiro, R., Macedo, D.R., Hughes, R.M., Kaufmann, P.R., Oliveira, L.G., Callisto, M., 2014. Importance of environmental factors for the richness and distribution of benthic macroinvertebrates in tropical headwater streams. Freshw. Sci. 33, 860–871.
Fohrer, N., Haverkamp, S., Eckhardt, K., Frede, H.-G., 2001. Hydrologic response to land use changes on the catchment scale. Phys. Chem Earth Part B Hydrol. Oceans Atmos. 26, 577–582.
Gao, X., Shi, Y., Giorgi, F., 2011. A high resolution simulation of climate change over China. Sci. China Earth Sci. 54, 462–472.
Grace, J.B., Pugesek, B.H., 1997. A structural equation model of plant species richness and its application to a coastal wetland. Am. Nat. 149, 436–460.
Grace, J.B., Schoolmaster Jr, D.R., Guntenspergen, G.R., Little, A.M., Mitchell, B.R., Miller, K.M., Schweiger, E.W., 2012. Guidelines for a graph-theoretic implementation of structural equation modeling. Ecosphere 3, 1–44.
Groisman, P.Y., Knight, R.W., Karl, T.R., 2001. Heavy precipitation and high streamflow in the contiguous United States: trends in the twentieth century. Bull. Am. Meteorol. Soc. 82, 219–246.
Guo, F., Kainz, M.J., Sheldon, F., Bunn, S.E., 2016a. Effects of light and nutrients on periphyton and the fatty acid composition and somatic growth of invertebrate grazers in subtropical streams. Oecologia 181, 449–462.
Guo, F., Kainz, M.J., Sheldon, F., Bunn, S.E., 2016b. The importance of high-quality algal food sources in stream food webs–current status and future perspectives. Freshw. Biol. 61, 815–831.
Gupta, S., Kim, H.W., 2008. Linking structural equation modeling to Bayesian networks: decision support for customer retention in virtual communities. Eur. J. Oper. Res. 190, 818–833.
Hansen, A.J., Neilson, R.P., Dale, V.H., Flather, C.H., Iverson, L.R., Currie, D.J., Shafer, S., Cook, R., Bartlein, P.J., 2001. Global change in forests: responses of species, communities, and biomes interactions between climate change and land use are projected to cause large shifts in biodiversity. Bioscience 51, 765–779.
Hershey, A.E., Lamberti, G.A., 1998. Stream macroinvertebrate communities. River Ecology and Management-Lessons from the Pacific Coastal Ecoregion. pp. 169–199.
Jhingran, V., Ahmad, S., Singh, A., 1989. Application of Shannon–Wiener index as a measure of pollution of river Ganga at Patna, Bihar, India. Curr. Sci. 717–720.
Jowett, I.G., Richardson, J., 1990. Microhabitat preferences of benthic invertebrates in a New Zealand river and the development of in-stream flow-habitat models for Deleatidium spp. N. Z. J. Mar. Freshw. Res. 24, 19–30.
Jowett, I.G., Richardson, J., Biggs, B.J., Hickey, C.W., Quinn, J.M., 1991. Microhabitat preferences of benthic invertebrates and the development of generalised Deleatidium spp. habitat suitability curves, applied to four New Zealand rivers. N. Z. J. Mar. Freshw. Res. 25, 187–199.
Karl, T.R., Trenberth, K.E., 2003. Modern global climate change. Science 302, 1719–1723.
Kath, J.M., Le Brocque, A.F., Maron, M., 2016. Using a Bayesian network model to assess ecological responses to hydrological factor interactions. Ecohydrology 9, 11–20.
Khan, S., Hanjra, M.A., Mu, J., 2009. Water management and crop production for food security in China: a review. Agric. Water Manag. 96, 349–360.
Klemm, D.J., Blocksom, K.A., Fulk, F.A., Herlihy, A.T., Hughes, R.M., Kaufmann, P.R., Peck, D.V., Stoddard, J.L., Thoeny, W.T., Griffith, M.B., 2003. Development and evaluation of a macroinvertebrate biotic integrity index (MBII) for regionally assessing Mid-Atlantic Highlands streams. Environ. Manag. 31, 0656–0669.
Kohonen, T., 1998. The self-organizing map. Neurocomputing 21, 1–6.
Kong, W., Meng, W., Zhang, Y., Gippel, C., Qu, X., 2013. A freshwater ecoregion delineation approach based on freshwater macroinvertebrate community features and spatial environmental data in Taizi River Basin, northeastern China. Ecol. Res. 28, 581–592.
Konrad, C., Brasher, A., May, J., 2008. Assessing streamflow characteristics as limiting factors on benthic invertebrate assemblages in streams across the western United States. Freshw. Biol. 53, 1983–1998.
Kragt, M.E., Newham, L.T., Bennett, J., Jakeman, A.J., 2011. An integrated approach to linking economic valuation and catchment modelling. Environ. Modell. Softw. 26, 92–102.
Kuemmerlen, M., Schmalz, B., Cai, Q., Haase, P., Fohrer, N., Jähnig, S.C., 2015. An attack on two fronts: predicting how changes in land use and climate affect the distribution of stream macroinvertebrates. Freshw. Biol. 60, 1443–1458.
Landuyt, D., Broekx, S., D'hondt, R., Engelen, G., Aertsens, J., Goethals, P.L., 2013. A review of Bayesian belief networks in ecosystem service modelling. Environ. Modell. Softw. 46, 1–11.
Lenat, D.R., Crawford, J.K., 1994. Effects of land use on water quality and aquatic biota of three North Carolina Piedmont streams. Hydrobiologia 294, 185–199.
Lessard, J.L., Hayes, D.B., 2003. Effects of elevated water temperature on fish and macroinvertebrate communities below small dams. River Res. Appl. 19, 721–732.
Li, X., Feng, J., Wellen, C., Wang, Y., 2017. A Bayesian approach of high impaired river reaches identification and total nitrogen load estimation in a sparsely monitored basin. Environ. Sci. Pollut. Res. 24, 987–996.
Liu, M., Hu, Y., Zhang, W., Zhu, J., Chen, H., Xi, F., 2011. Application of land-use change model in guiding regional planning: a case study in Hun-Taizi River Watershed, Northeast China. Chin. Geogr. Sci. 21, 609–618.
Malaeb, Z.A., Summers, J.K., Pugesek, B.H., 2000. Using structural equation modeling to investigate relationships among ecological variables. Environ. Ecol. Stat. 7, 93–111.
Marcot, B.G., Holthausen, R.S., Raphael, M.G., Rowland, M.M., Wisdom, M.J., 2001. Using Bayesian belief networks to evaluate fish and wildlife population viability under land management alternatives from an environmental impact statement. For. Ecol. Manag. 153, 29–42.
Marcot, B.G., Steventon, J.D., Sutherland, G.D., McCann, R.K., 2006. Guidelines for developing and updating Bayesian belief networks applied to ecological modeling and conservation. Can. J. For. Res. 36, 3063–3074.
McCann, R.K., Marcot, B.G., Ellis, R., 2006. Bayesian belief networks: applications in ecology and natural resource management. Can. J. For. Res. 36, 3053–3062.
Merritt, R.W., Cummins, K.W., 1996. An Introduction to the Aquatic Insects of North America. Kendall Hunt.
Meyer, W.B., Turner, B.L., 1992. Human population growth and global land-use/cover change. Annu. Rev. Ecol. Syst. 23, 39–61.
Min, S.-K., Zhang, X., Zwiers, F.W., Hegerl, G.C., 2011. Human contribution to more-intense precipitation extremes. Nature 470, 378–381.
Murphy, J., Riley, J.P., 1962. A modified single solution method for the determination of phosphate in natural waters. Anal. Chim. Acta 27, 31–36.
Narangarvuu, D., Oyunbileg, J., Yang, P.-S., Boldgiv, B., 2015. Distribution of Ephemeroptera, Plecoptera, and Trichoptera assemblages in relation to environmental variables in headwater streams of Mongolia. Environ. Earth Sci. 73, 835–847.
Neal, R.M., Hinton, G.E., 1998. A View of the EM Algorithm That Justifies Incremental, Sparse, and Other Variants, Learning in Graphical Models. Springer, pp. 355–368.
Ngo-Duc, T., Oki, T., Kanae, S., 2007. A variable streamflow velocity method for global river routing model: model description and preliminary results. Hydrol. Earth Syst. Sci. Discuss. 4, 4389–4414.
Novotny, E.V., Stefan, H.G., 2007. Stream flow in Minnesota: indicator of climate change. J. Hydrol. 334, 319–333.
Ometo, J.P.H., Martinelli, L.A., Ballester, M.V., Gessner, A., Krusche, A.V., Victoria, R.L., Williams, M., 2000. Effects of land use on water chemistry and macroinvertebrates in two streams of the Piracicaba river basin, south-east Brazil. Freshw. Bio. 44, 327–337.
Parmesan, C., Yohe, G., 2003. A globally coherent fingerprint of climate change impacts across natural systems. Nature 421, 37.
Paul, M.J., Meyer, J.L., 2001. Streams in the urban landscape. Annu. Rev. Ecol. Syst. 32, 333–365.
Peterson, R., Eeckhaute, L., 1992. Distributions of Ephemeroptera, Plecoptera, and Trichoptera of three maritime catchments differing in pH. Freshw. Biol. 27, 65–78.
Pond, G.J., 2012. Biodiversity loss in appalachian headwater streams (Kentucky, USA): plecoptera and trichoptera communities. Hydrobiologia 679, 97–117.
Rempel, L.L., Richardson, J.S., Healey, M.C., 2000. Macroinvertebrate community structure along gradients of hydraulic and sedimentary conditions in a large gravel-bed river. Freshw. Biol. 45, 57–73.
Roy, A., Rosemond, A., Paul, M., Leigh, D., Wallace, J., 2003. Stream macroinvertebrate response to catchment urbanisation (Georgia, USA). Freshw. Biol. 48, 329–346.
Sangiorgio, F., Quintino, V., Rosati, I., Rodrigues, A.M., Pinna, M., Basset, A., 2014. Macrofauna in Mediterranean and Black Sea transitional aquatic ecosystems: a comparative study of the benthic populations sampled by box corer and leaf bags. Ecol. Indic. 38, 159–169.
Schmitt, R., Siegloch, A.E., da Silva, A.L.L., Lisboa, L.K., Petrucio, M.M., 2016. Temporal variation in the Ephemeroptera, Plecoptera and Trichoptera community in response to environmental drivers in a subtropical stream. J. Insect Biodivers. 4, 1–12.
Sponseller, R., Benfield, E., Valett, H., 2001. Relationships between land use, spatial scale and stream macroinvertebrate communities. Freshw. Biol. 46, 1409–1424.
Swait, J., 1994. A structural equation model of latent segmentation and product choice for cross-sectional revealed preference choice data. J. Retail. Consum. Serv. 1, 77–89.
Tanaka, J.S., Huba, G.J., 1985. A fit index for covariance structure models under arbitrary GLS estimation. Br. J. Math. Stat. Psychol. 38, 197–201.
Tejerina-Garro, F.L., Maldonado, M., Ibañez, C., Pont, D., Roset, N., Oberdorff, T., 2005. Effects of natural and anthropogenic environmental changes on riverine fish assemblages: a framework for ecological assessment of rivers. Braz. Arch. Biol. Technol. 48, 91–108.
Tong, S.T., Chen, W., 2002. Modeling the relationship between land use and surface water quality. J. Environ. Manag. 66, 377–393.
Vörösmarty, C.J., McIntyre, P.B., Gessner, M.O., Dudgeon, D., Prusevich, A., Green, P., Glidden, S., Bunn, S.E., Sullivan, C.A., Reidy Liermann, C., 2010. Global threats to human water security and river biodiversity. Nature 467, 555.
Walsh, C.J., 2006. Biological indicators of stream health using macroinvertebrate assemblage composition: a comparison of sensitivity to an urban gradient. Mar. Freshw. Res. 57, 37–47.
Wan, J., Bu, H., Zhang, Y., Meng, W., 2013. Classification of rivers based on water quality assessment using factor analysis in Taizi River basin, northeast China. Environ. Earth Sci. 69, 909–919.
Watanabe, K., Yaegashi, S., Tomozawa, H., Koshimura, S., Omura, T., 2014. Effects on river macroinvertebrate communities of tsunami propagation after the 2011 Great East Japan Earthquake. Freshw. Biol. 59, 1474–1483.
Wiggins, G.B., 2015. Larvae of the North American Caddisfly Genera (Trichoptera). University of Toronto Press.

Wilkins, P.M., Cao, Y., Heske, E.J., Levengood, J.M., 2015. Influence of a forest preserve on aquatic macroinvertebrates, habitat quality, and water quality in an urban stream. Urban Ecosyst. 18, 989–1006.

Witten, I.H., Frank, E., 2005. Data Mining: Practical Machine Learning Tools and Techniques. Morgan Kaufmann.

Zhang, Y., Guo, F., Meng, W., Wang, X.Q., 2009. Water quality assessment and source identification of Daliao River Basin using multivariate statistical methods. Environ. Monit. Assess. 152, 105–121.

Zhang, Y., Zhao, R., Kong, W., Geng, S., Bentsen, C.N., Qu, X., 2013. Relationships between macroinvertebrate communities and land use types within different riparian widths in three headwater streams of Taizi River, China. J. Freshw. Ecol. 28, 307–328.